Identity Verification and Fraud Detection During Online Exams with a Privacy Compliant Biometric System

M. A. Haytom[1,2], C. Rosenberger[1], C. Charrier[1], C. Zhu[2] and C. Regnier[2]
[1]*Normandie Univ., UNICAEN, ENSICAEN, CNRS, GREYC, 14000 Caen, France*
[2]*TestWe, 75003 Paris, France*

Keywords: Personal Data, Biometric Authentication, Privacy Protection, Machine Learning.

Abstract: Distant learning is an alternative solution to education when the learner is far from the school or cannot attend courses for professional or medical reasons. The main objective of this work is to design a smart application of remote exams, using a multibiometric system combining face with deep learning and keystroke dynamics to verify the identity of the learner. Privacy protection is consider in this work as an important issue because many personal data are processed in the proposed solution. We consider in this paper experiments under real-life conditions to identify abnormal behaviours with confidence indicators. We show the system ability to make the correct decision while preserving learner's privacy.

1 INTRODUCTION

Today, with the rapid growth of online courses and exams, more and more students or professionals are applying for distance learning. Certainly, distance learning offers several advantages: the opportunity to take a training course from a distant establishment, arrange teaching for the personal and professional life of the learner, have great organizational flexibility and reduce accommodation and transportation costs. It allows to have the same training level obtained in real classrooms. Yet, secure tools must be used to safely access the exams. It is, thus, essential to control the student access and environment during an exam. Online assessment and supervised tests address the issue of student verification and the environment in which access to unauthorized documents and resources is a major problem.

Data processing and environmental control must respect the principles defined by the General Regulation on the protection of personal information (Carey, 2018). The information often collected in a regulatory context can make the verification of a person a safe and secure step. Higher education encourages to fight cheating more effectively, schools are so very interested in the quality and integrity of online proctoring. There is strong evidence that cheating has increased in classrooms as it has been shown that 70% of students cheat during their schooling (Chauvel,). Various possible attacks that threaten the privacy of individuals can also be observed. The student may be faced with an identity theft, whether hacking his/her account, or attempting to perform fraudulent actions. It is therefore mandatory to count and identify them to define the necessary requirements for the management of exams and the protection of personal data.

The need for reliable user authentication solutions has increased with growing security concerns and rapid advances in networking and communication. Biometrics refers precisely to a computer technology that stimulates the physical presence of an individual, it conceives its identity and its movements as sources of dangers and risks. It is described as the science of recognizing an individual based on their physical or behavioral characteristics. This technology is an excellent candidate for verifying the identity of an individual. Biometrics are divided into two categories: 1) physical and 2) behavioural. Physical analysis may include the geometry of the hand, the retina, the characteristics of the iris and so on. Behavioural analysis helps to verify a person's identity by examining measurable activity, such as gesture recognition, keystroke dynamics and so on.

There are two main contributions in this work: 1) proposal of an original authentication solution based on facial recognition and keystroke dynamics for online assessments and the integration of a security system to detect fraud during online exams and 2) per-

Haytom, M., Rosenberger, C., Charrier, C., Zhu, C. and Regnier, C.
Identity Verification and Fraud Detection During Online Exams with a Privacy Compliant Biometric System.
DOI: 10.5220/0009874104510458
In *Proceedings of the 17th International Joint Conference on e-Business and Telecommunications (ICETE 2020) - SECRYPT*, pages 451-458
ISBN: 978-989-758-446-6

forming experiments in real conditions to validate the proposed approach (Fig. 1). The paper is organized as follows. In section 2, we present the related works on the identity verification and fraud detection solutions during online examinations. We present the proposed system in section 3. Section 4 is dedicated to experimental results. We finally conclude in section 5 and give some perspectives of this work.

2 RELATED WORKS

In the literature, we can only find few studies that deal with remote exam management. Duchatelle et al. conducted a real-life remote monitoring experiment at the University of Caen Normandie (P. Beust, 2016) allowing students in different countries to take this exam. According to the obtained results, several students have accepted to participate in the experiment. Half of the candidates go as far as creating an account on the platform of the organism that manages the exams. As part of this experiment, 31 participants completed the test and almost a third of the candidates could not pass the test due to technical problems. This experience remains among the first online distance exams using biometric-based authentication. Recently, authors in (Arnautovski, 2019) proposed a face recognition system for online examinations. Yet, the proposed system provides no privacy protection and has not been tested in real conditions.

2.1 Keystroke Dynamics

Many applications nowadays require personal data about the user, where the proof of identity is used to provide a secure service and to guarantee that a person is authorized to access an online multimedia service. In addition, it is necessary to convince users that the proposed solution is credible and that the information is processed reliably. Without a doubt, keystroke dynamics could play a necessary role to improve the identity verification of individuals. This biometric solution has a very high degree of acceptability when compared to other modalities such as iris or gesture recognition.

Keystroke dynamics represents the analysis of keyboard interactions (for example the flight time between two pressed keys), recorded during the use of a laptop by the user. The collected information describe the time when the key is pressed and the time when the key is released when the learner uses his keyboard. Researchers have carried out numerous studies in this field, it is possible to explore these information for the authentication of an individual or profiling purpose in order to recognize the gender or the age category according to the password or the edited text with a recognition rate close to 75% (Giot et al., 2012; Idrus et al., 2013). In fact, it is possible to combine keystroke data with other biometric modality in order to have a more secure verification system. In the first place, we have integrated keystroke dynamics into our remote exam management system as a basic modality to increase the level of security and to have a robust identity verification process.

2.2 Facial Recognition

Facial recognition systems have become an useful tool in many fields of application. It remains an acceptable solution for many people because it is less expensive and easier to use. Facial recognition systems are well-known computer vision applications whose purpose is to verify or identify a person. There are many approaches for facial recognition in the literature, such as: morphological elastic graph matching (Kotropoulos et al., 2000), scale-invariant feature transform (Bicego et al., 2006) and so on.

However, testing in a non-controlled environment (variable lighting, variable orientation of the person towards the camera, moving person, etc.) remains problematic, since the verification of users is carried out with high uncertainty especially during the acquisition phase. Certainly, many other methods to verify the identity of a person such as Deep Learning based approach exist (Liu et al., 2015). To address the critical environmental issue, the research activity focused on the analysis of human perception in order to mathematically describe the processes that lead to the recognition of actions, in order to mimic to the recognition process of faces and objects by human beings. Today, there are several works and scientific publications based on artificial intelligence for user verification and identification (Liu et al., 2015; Parkhi et al., 2015).

Face recognition is a constantly evolving field, continually improving. Identity verification through deep learning has recently received a lot of attention from the research and industry community and is starting to be applied in a variety of areas, mainly for security. We can find a study that handles the authentication based on a Convolution Neural Network (Parkhi et al., 2015), researchers compute VGG-Face CNN descriptors using CNN implementation based on the VGG-Very-Deep-16 CNN architecture and they evaluate it on a large databases (Labeled Faces in the Wild and the YouTube Faces). Deep learning remains an effective technique and works much better than many facial recognition methods.

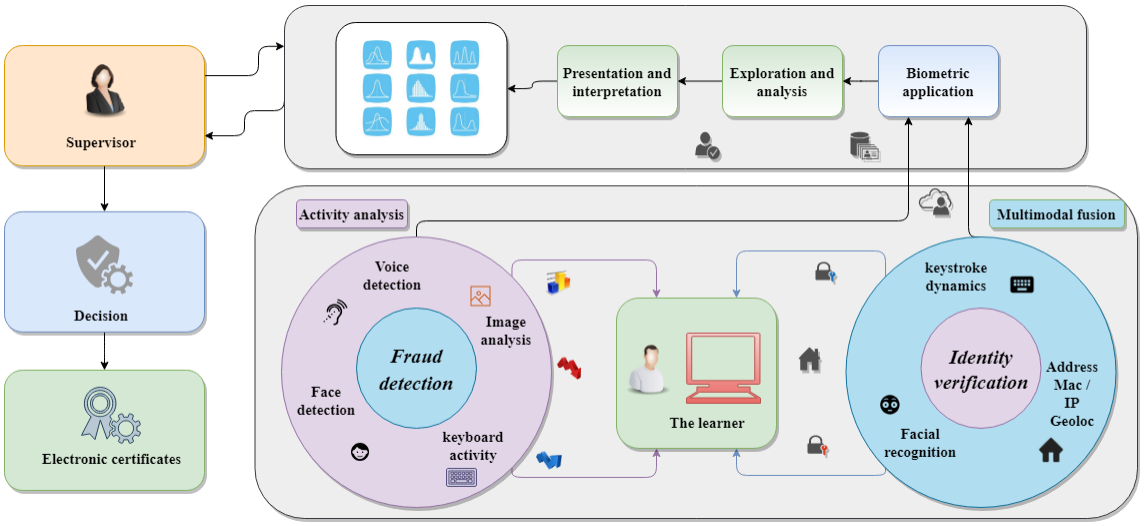

Figure 1: Multimodal biometric system with an automated monitoring.

The combination of keystroke dynamics and facial recognition has also been studied. Unlike (Gupta et al., 2015) and (Giot et al., 2010), our approach is based on a deep learning solution for faces and then the use of a different protocol of keystroke dynamics for free text field. In addition, the use of biometric and non-biometric data to form a model allowing identity verification and fraud detection in a remote examination.

3 PROPOSED APPROACH

In this paper, we propose the integration of both facial recognition solution and keystroke into the remote examination management system as a multimodal solution to increase the security level of the application. In fact, for the faces, we improved the pre-processing phase to have a better presentation for data analysis before using deep learning process to extract the characteristics of a learner. In fact, our objective is to set up a combination of several existing techniques in a remote examination management system to have a reliable and efficient solution. The biometric signatures are part of a lot of so-called sensitive data. So, we implement security measures and data protection processes by obtaining an encrypted template in the form of a secure code. The biohashing algorithm has been used to ensure the confidentiality of the data exchanged since the storage of biometric templates is a major problem.

Moreover, we propose the use of a fraud detection model to avoid frauds and identity theft problems during remote exams. In this work, we conducted an experiment in real conditions in order to validate the proposed solution. Several fraud scenarii have been considered; direct communication with a person nearby, the use of unauthorized documents, the use of a second screen smartphone and/or tablet to connect to the internet or sound transmission as sound effects with fingers or hands or even leaving the room to have answers, etc. The steps of the proposal solution are:

Authentication
1. Keystroke dynamics process
2. Face recognition process
3. Compute for each feature vector its Biocode

Fraud detection
1. Perform a continuous verification
2. Create fraud detection model

3.1 Identity Verification

3.1.1 Keystroke Dynamics Process

For the first biometric modality, we carry out a continuous verification throughout the exam session for each learner. First of all, when the user writes a text using his keyboard, each word represents a signature. Words with more than two letters are processed and analyzed. From several time sequences, we extract the learner's template where: 1) the time interval between pressed and released key, 2) the time interval between two consecutive pressures, 3) the time interval between a release and the pressing of the next key and 4) the overall duration of a series of characters.

Unlike several state of the art studies, the learner's hesitation time was taken into account with the erro-

neous data. The hesitation times to write a text reveal a unique data related to the person and his real way of using his keyboard, which is not the case for passwords. For each keystroke, the time intervals are recorded and concatenated with the duration to constitute a vector of characteristics.

Keystroke dynamics can be affected by a number of variables (how well the learner prepared for different exams, and how familiar he/she is with the topic of each question within the same exam, etc), to avoid this, during the enrollment the user is invited to write a text of a page, the hesitation time is taken into account with a delay of 3 seconds maximum. Depending on the security requirements of the application, the KD is used in this context only for texts at least one page long.

3.1.2 Facial Recognition Process

Face recognition systems involve two important phases, the enrollment and the authentication. The first step concerns the user enrollment: enrollment means the acquisition of the biometric input data, the detection of region of interest and the extraction of features to define a template of each genuine user. The genuine model is stored as a reference after applying some quality requirements (Saad et al., 2012). For face recognition based systems, the concerned biometric data is the face image, and the extracted features are stored in the user device memory.

To achieve the enrollment phase in a reliable manner on a massive scale, a familiarization session was planned to get started with the software one day before the exam. During this simulation the learner is invited to capture his image, this image will be used in order to make a verification during the real session of the exam.

On the first hand, to authenticate users, we start by capturing images using the student laptop camera. A pre-processing step has been used to enhance the input data. Once the face image has been acquired, we use a face detection tool to crop and normalize the faces. Next, we extract features using the Convolutional Neural Network (CNN) model (Parkhi et al., 2015) to obtain a template representing the signature of the user. With the biometric model, students can access online courses and exams depending on the threshold used for the verification process. First, we compare the reference model with the signature acquired during authentication and we compute a score distance.

Quality Evaluation
The no-reference image quality assessment algorithm BLIINDS2 (Saad et al., 2012) has been used to evaluate the quality of the image, using a natural scene statistical model of discrete cosine transform coefficients. The approach relies on a simple Bayesian inference model to predict image quality scores based on certain extracted features. During the pre-processing step, if the image quality metric is lower than the threshold, the acquired image is not saved, otherwise the image is recorded.

Adjusting Intensity
The values of the intensity of the image have been adjusted to improve the contrast of the output image by saturating the higher side and the lower side by 1 percent of all pixel values.

Face Detection
In this part, Viola-Jones Cascade Object Detector has been used to detect the faces (Viola and Jones, 2001). Indeed, the calculation is based on the integration of the pixels of the image as well as the use of pseudo-Haar features. After getting the size and the position or the spatial coordinates of the region of interest. The output is cropped including all the pixels of the image inside the rectangle. The image is resized by specifying the number of rows and columns to get a matrix of two dimensions (resolution 224×224 pixels).

VGG Face - Convolution Neural Network
CNN architectures were mainly used for character recognition tasks. They are proven very effective in areas such as image recognition and classification (Parkhi et al., 2015). The model was formed on 2.6 million images over more than 2,600 people. The accuracy of the system has been evaluated at 98.95%, compared to DeepFace and FaceNet system. To extract the feature vector, VGG pre-training model has been applied to the input image (resolution 224×224 pixels), the output data represents the user pattern or the unique signature of the learner.

3.1.3 Privacy Protection

The general principle of BioHashing is to generate a binary code called Bio-Code, which will be used during the enrollment phase and to verify the identity of the learner. It can be obtained from the biometric data (the template of the user) and a random number. In the literature, we can find many studies using Biohashing for many biometrics modalities such as fingerprints (Belguechi et al., 2010; Belguechi et al., 2012), keystroke dynamics (Migdal and Rosenberger, 2018) and so on.

In this work, we use the BioHashing algorithm in order to protect the VGG face vector and keystroke

signature computed from an user for privacy reasons. Face images and keystroke data are not stored for the same reasons. This algorithm consists firstly in projecting the native biometric data on an orthogonal basis generated from the random number. The resulting dimension corresponds at most to the dimension of the primary representation of the biometric data. The goal of this step is to hide the biometric data in a part of the multidimensional space. The use of an orthogonal base ensures the conservation of the similarity relations between the templates (BioCodes), while kepping good recognition result. The second step aims to quantify this result using a determined threshold. In our study, for face the size of the generated biocode is 512 bits. The choice of the threshold and the size of the bioCode guarantees the non-inversion of the process. In addition, this two-factor process (FingerCode and Seed-key) ensures that it is not possible to retrieve the native biometric data given the BioCode.

3.2 Fraud Detection

In our study, we use face detection, analysis of events using images and facial recognition by performing a calculation on a four seconds duration window. In fact several parameters were used: the keyboard activity, the keystroke dynamics and the acoustic signal. On the other hand, we perform an analysis for each type of information, then we use statistical solutions such as the histogram in order to distinguish honest learners and learners who cheat during the exam. For example, for face detection, the histogram allows to represent the distribution of the numbers of detection continuously until the end of the test. The first step consists in comparing the data acquired from a person authorized to take the examination with and without fraud attempts to the information collected from a person not authorized to take the examination. When the amplitude of the signal is higher than the threshold at a given moment, we consider that the user cheated. To set the threshold, some machine learning tools have been used in order to create a fraud detection model by automatically setting the threshold value.

The threshold could have a significant impact on performance. During this study we note that, the observations are well separated between a person who cheats and an honest learner. In fact, the use of several characteristics (data from different sensors) drastically reduce the impact of the threshold on performance.

We trust the output provided by the camera, even if in a way, a user can rely on techniques similar to those provided by malware to produce a fake video stream. In fact, the basic exam management solution blocks access to internet during the examination with a secure virtual environment, and the learner can't manipulate other applications before sending the copy of the exam and closing the application.

3.2.1 Fraud Model

Here, we propose to use a machine learning method to detect the fraud automatically and creating a fraud detection model to manage online proctoring. The dataset for training contains normal cases and frauds. The model was created using four supervised classification algorithms: 1) the Decision Tree, 2) the Logistic regression, 2) the k-nearest neighbors scheme and 4) the Naive Bayes classifier after pre-processing and re-sampling to obtain the user indicators. Indicators were obtained after launching a continuous analysis for each user.
In fact, a set of biometrics and non biometrics data (25 learners) have been used to train the model, each sample contains; 1) face detection analysis, 2) the keyboard event (by counting the number of keys), 3) the distributions of image analysis, 4) three distributions based on the amplitude of the sound signal by varying the amplitude thresholds 5) the keystroke dynamics and 6) the feature vectors of the deep parameter of facial recognition every four seconds, each signature of face was obtained after running the process of deep learning. To form our detection model, data from various scenarios were used. We start the training by using a class with 703 characteristics, for each user 150 instances, making a total of 3750 instances. The model was trained to count cheating attempts and identify frauds automatically.

4 EXPERIMENTAL RESULTS

4.1 Experimental Protocols

4.1.1 Experiment 1

In this work, biometric data (face images) from 30 users working in the GREYC Laboratory and TestWe Company were collected using a laptop computer. This number of participants is not so high, but is in the same range as published studies (see section 2). Biometric information is collected from individuals via an integrated camera of a "ThinkPad S440" computer. Intel Core i7-4510U computer with a CPU of 2.00 and 2.60 [GHz] was used during the examination.

This first experiment remains a preliminary study to have the feelings of the users, evaluate the pro-

posed method and the system ability to make the right decision. Each user is invited to take the exam remotely with a graphical user interface (QT application). More precisely, the sample is recorded and it is the same for each user. Experiments were carried out with two different protocols: the first protocol consists of passing the test in an uncontrolled environment for ten participants and the second protocol twenty people have passed the test with the same device in a controlled environment. The constraints of the uncontrolled environment are: the variation of light in the examination room, the face pose and the presence of posters and images in the background of the learners. For this preliminary study, the database contains a number of 30 people and represents a class of examination, for each participant 720 images are collected making a total of 21600 samples.

4.1.2 Experiment 2

We have created an application with a text editor to collect keystroke data. The application allows us to collect the flight time between each pressed key as well as the release time of the keys. Two sessions were carried out (the same number of participants of the previous experience); during the first session, participants were asked to write one page of text and then write ten sentences before sending the data. The second session consists of launching a second application (keylogger) for a period of one day in order to collect as much data as possible. The number of words recorded for each learner varies with an average of 1200. We take the minimum value (150 samples) in order to process the data of all participants. A standby button is used when a participant enters personal information (password, PIN, etc.).

4.1.3 Experiment 3

A third experiment in real conditions was carried out with ENSICAEN students to simulate fraud attempts in a remote examination. Several fraud scenarios have been considered using unauthorized supports (smartphone, blackboard, second computer, etc.) or the help of a friend nearby (direct or indirect communication) or even leaving the room to seek answers. The overall duration is 10 minutes for each simulation. User interface application (QT C++) has been used to collect the data, the information collected are the audio signal, the sequence of images then the flight time between each used key.

Table 1: Evaluation in an uncontrolled and in a controlled environment (Equal Error Rate).

Template	uncontrolled env		controlled env	
Ref (aver)	FaceC	BioC	FaceC	BioC
10	0.531	0.292	0,0099	0.001
20	0.529	0.26	0,0088	0.002
30	0.444	0.24	0,0069	0.001
40	0.35	0.254	0,0057	0.001
50	0.316	0.282	0.0065	0.001

4.2 Performance Evaluation

We consider that the conditions related to the sensors are fulfilled. Some common error rates in biometrics will be used such as Equal Error Rate (EER) with computes the system configuration to have a tradeoff between the FAR (false acceptance rate) and the FRR (false rejection rate) values.

4.2.1 Face Recognition

We consider the database n1 composed of 720 images (resolution 224×224 pixels) for 10 people. The faceCode of each user is generated according to the method presented in section 3.1, with deep learning. Once this is done, the biohashing has been used, 720 faceCodes are available for each person, which means 7200 faceCodes with a length of 512 bits. After random projection and quantization, 7200 BioCodes are generated. For each person, the average of several FaceCodes is kept as a reference, the other FaceCodes are used to evaluate the performance of the system.

For the first scenario, the reference FaceCodes was chosen after calculating the average of the first ten FaceCodes. Consecutively, the number of samples increased from 10 to 50 with a step of 10. On the other hand, for the bioCodes, the reference templates were chosen after calculating the average in a consecutive way. The average of several BioCodes is used as a reference, the other BioCodes are also used for the comparison. The BLIINDS2 quality assessment algorithm (Saad et al., 2012) was not used to increase the false acceptance rate during the pre-processing phase. The EERs obtained (Table 1) are apparently far from being the best (with errors related to non detection) compared to the results of the literature. However, these values allow us to define a basic system performance in an uncontrolled environment.

In the second scenario, the process of evaluating the image quality was used, which reduces the number of false acceptances. In the table (Table 1), we see that in a controlled environment, the error rates decrease. The EER value of Facecode is greater than the EER value of the BioCode in both scenarios. Indeed,

Table 2: Evaluation of two sessions (EER).

ref template	BioHash		Native	
(average 1-10)	min	max	min	max
Keylogger	0.019	0.16	0.002	0.004
Normal sess	0.045	0.079	0.006	0.022

Table 3: The results of the fusion of faceCode scores and secure keystroke signature (EER).

ref template	BioCode EER	KD EER	merged score
10	0.0012	0,0813	0,0813
20	0.0018	0,0738	0,07378
30	0.0009	0,0722	0,07225
40	0.0009	0,0859	0,08592
50	0.00095	0,0880	0,08803

without the constraints related to the environment and the quality of the input data, after using bioHashing the error rates are close to zero. However, it must at least the average of the first forty FaceCodes to find an optimal reference template.

4.2.2 Keystroke Dynamics

The same principle to secure the face data was applied to protect the keystroke dynamics templates. Based on the results of our experiment, we see that the EER values vary between 0.2% and 2% for native data, and between 4 % and 22 % with bioHashing (table 2, keylogger and normal session). This is due to the fact that the size of the characteristic vector is very small. On the other hand, to secure more and more the biometric data, we increase the size of the keycode (the keystroke signature) before applying bioHashing by adding relevant information related to the learner. We note that the test based on the keylogger gives lower error rates than the normal session since this test allows participants to use their keyboard in a natural way.

4.2.3 Multimodal Fusion

The fusion during the generation of scores induces several problems such as non-homogeneous scores, various distributions and vectors of different sizes. In our study, we obtain an error rate of 8% with the average of the first 10 templates (Table 3, the results of the merger of scores), a considerably reduced error percentage compared to fusion in the extraction of characteristics. Indeed, for the reference template, the average of 10 to 50 templates was used for the signature of the face and the keystroke dynamics.

4.2.4 Fraud Detection Model

The idea here is to highlight the classification tools to create a model that will be used to detect fraud during the online exam. To detect fraud, multiple information of 25 participants from different sensors were used to generate the model. The target is: ”fraud attempt” or ”honest learner”. We chose all the biometric and non-biometric data available from these authors that were sufficiently numerous (with a 10 minutes duration exam), which were sufficient to carry out an analysis and determine the fraud attempts.

Table 4: Evaluation results (Cross validation, k = 20).

Model	AUC	CA	Precis	Recall
kNN	0.999	0.995	0.995	0.995
Tree	1	1	1	1
NB	0.987	0.962	0.962	0.962
LR	1	1	1	1

In order to establish the automated detection system, a set of data was used for each user (set of characteristics: face detection, facial recognition, video activity, keystroke dynamics, sound signal and keyboard events). We use different supervised classification approaches to build the model (Decision Tree (DT), logistic Regression (LR), k-nearest neighbors (kNN) and Naive Bayes (NB)).

A confusion matrix has been used to evaluate the proposed fraud detection method. After setting the output variable (”honest learner” and ”fraud attempt”) and comparing the detection accuracy of different machine learning algorithms by adjusting the parameter of the reliability estimation method (cross validation). we obtain a precision close to 100% with the Logistic Regression and the Decision Tree using all the characteristics, (Table 4). These two methods allow to identify almost all fraud attempts (detecting 2100 with the DT and 2099 fraud attempts with the LR, see Tables 5). After carrying out the experiments, the DT and RL learning methods provide the best performance. The DT and the RL remain a relevant solution for the classification and the detection of fraud in a remote exam. To conclude, we can integrate these two learning solutions into our biometric system to detect an anomaly or the unusual behaviour using a fraud detection model in order to improve online assessment.

Table 5: Confusion matrix - Tree and LR.

	Tree (predicted)		LR (predicted)	
	honest	fraud	honest	fraud
honest (act)	1650	0	1650	0
fraud (act)	0	2100	1	2099

5 CONCLUSION AND PERSPECTIVES

This work addresses the general problem of the protection of the privacy of a user during a remote examination. It seeks to define the different forms it can take fraud, as well as the expected properties of a secure biometric anti-cheat system that respects the privacy of learners. To conclude, the integration of biometrics into distance learning systems will help teachers to effectively control student authentication, course tracking, provide certificates in an automated manner, analyse student behaviour during exams, the validation of the certificates of success as well as the detection of fraud with a high level of accuracy. The proposed solution is effective against identity theft and fraud attempts. Indeed, this system is able to detect fraud automatically, it respects the privacy and confidentiality of the data exchanged and solves an important part of a major problem.

Concerning perspectives, we intend to evaluate the system performance after merging several modalities (face, keystroke dynamics, gaze, gestures and so on) then, create a fraud model using a larger database. In addition, the different scores can be combined to generate a confidence index on the identity of the learner during the examination. After all, other solutions could be included to detect spoofing attacks.

REFERENCES

Arnautovski, L. (2019). Face recognition technology in the exam identity authentication system-implementation concept. *Proceedings of Papers*, page 50.

Belguechi, R., Cherrier, E., and Rosenberger, C. (2012). Texture based fingerprint biohashing: Attacks and robustness. In *2012 5th IAPR International Conference on Biometrics (ICB)*, pages 196–201. IEEE.

Belguechi, R., Rosenberger, C., and Ait-Aoudia, S. (2010). Biohashing for securing minutiae template. In *2010 20th International Conference on Pattern Recognition*, pages 1168–1171. IEEE.

Bicego, M., Lagorio, A., Grosso, E., and Tistarelli, M. (2006). On the use of sift features for face authentication. In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, pages 35–35. IEEE.

Carey, P. (2018). *Data protection: a practical guide to UK and EU law*. Oxford University Press, Inc.

Chauvel, C. 70% des élèves trichent pendant leur scolarité.

Giot, R., Hemery, B., and Rosenberger, C. (2010). Low cost and usable multimodal biometric system based on keystroke dynamics and 2d face recognition. In *2010 20th International Conference on Pattern Recognition*, pages 1128–1131. IEEE.

Giot, R., Rosenberger, C., and Dorizzi, B. (2012). Reconnaissance du genre par analyse de dynamique de frappe au clavier sur texte libre.

Gupta, A., Khanna, A., Jagetia, A., Sharma, D., Alekh, S., and Choudhary, V. (2015). Combining keystroke dynamics and face recognition for user verification. In *2015 IEEE 18th International Conference on Computational Science and Engineering*, pages 294–299. IEEE.

Idrus, S. Z. S., Cherrier, E., Rosenberger, C., and Bours, P. (2013). Soft biometrics for keystroke dynamics. In *International Conference Image Analysis and Recognition*, pages 11–18. Springer.

Kotropoulos, C., Tefas, A., and Pitas, I. (2000). Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions. *Pattern Recognition*, 33(12):1935–1947.

Liu, Z., Luo, P., Wang, X., and Tang, X. (2015). Deep learning face attributes in the wild. In *Proceedings of the IEEE international conference on computer vision*, pages 3730–3738.

Migdal, D. and Rosenberger, C. (2018). Towards a personal identity code respecting privacy.

P. Beust, V. Cauchard, I. D. (2016). Enseigner autrement.

Parkhi, O. M., Vedaldi, A., and Zisserman, A. (2015). Deep face recognition.

Saad, M., Charrier, C., and Bovik, A. C. (2012). Evaluation de la qualité des images sans référence par modélisation statistique. COmpression et REprésentation des Signaux Audiovisuels (CORESA).

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, volume 1, pages I–I. IEEE.